Virology 433 (2012) 282–295

Contents lists available at SciVerse ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Selection of variant viruses during replication and transmission of H7N1 viruses in chickens and turkeys

Munir Iqbal [a,*], Steve C. Essen [b], Haixia Xiao [c], Sharon M. Brookes [b], Ian H. Brown [b], John W. McCauley [c]

[a] Avian Viral Diseases Programme, Institute for Animal Health, Compton Laboratory, Compton, Newbury, Berkshire RG20 7NN, UK
[b] Avian Virology, Animal Health and Veterinary Laboratories Agency-Weybridge, Addlestone, Surrey KT15 3NB, UK
[c] Division of Virology, MRC National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

ARTICLE INFO

Article history:
Received 6 April 2012
Returned to author for revisions
25 May 2012
Accepted 1 August 2012
Available online 1 September 2012

Keywords:
Influenza
Glycosylation
Haemagglutinin
In vivo infection
Poultry

ABSTRACT

The influence of different glycosylation patterns of the haemagglutinin glycoprotein of H7N1 avian influenza viruses on virus replication *in vivo* was examined. Experimental infection of chickens and turkeys was carried out with H7N1 avian influenza viruses with alternative sites of glycosylation in the haemagglutinin and infected birds were sampled daily by swabbing the buccal and cloacal cavities. cDNAs of the HA1 coding region of the HA gene were prepared from the swabs and cloned into plasmids. Sequencing multiple plasmids made from individual swabs taken over the period of virus shedding showed that viruses with specific patterns of glycosylation near the receptor binding site were stable when birds were infected with a single variant, but when presented with a mixed population of viruses encoding differing patterns of glycosylation a specific variant was rapidly selected in the infected host.

Introduction

Avian influenza (AI) viruses naturally circulate in wild aquatic birds as viruses of low pathogenicity (LP) for poultry but transmission to other avian species can result in more serious sporadic or sustained infections. Transmission of highly pathogenic (HP) AI viruses to domestic poultry or mutation of LP viruses to HP after infection of poultry results in serious disease in which mortality approaching 100% can be observed. However, LP avian influenza viruses (LPAI) can cause significant disease when exacerbated by other infections or unfavourable environmental conditions (Brown et al., 2006; Capua and Alexander, 2009).

Sixteen haemagglutinin (HA) and nine neuraminidase (NA) subtypes of the virus surface glycoproteins have been recognised in viruses collected from birds in the order Anseriformes or in the order Charariiformes; a sub-set of which circulate in mammals, but recently a new sub-type (H17N10) has been reported to be present exclusively in the mammalian order Chiroptera (Olsen et al., 2006; Tong et al., 2012; Webster et al., 1992). Transmission of viruses from water birds to a new host species does not always result in a successful persistence and spread of the virus within the population, and replication in a new species will impose pressures on the virus to adapt to the new host. Transmission from wild bird to domestic poultry species is seen frequently (Alexander and Brown, 2000; Webster et al., 1997) and can cause mild to severe disease (Alexander, 2007; Xu et al., 2007). Infection by HPAI and LPAI viruses in domesticated poultry represents a substantial threat to poultry production and can have a serious economic impact. Avian influenza viruses also can pose a risk to public health through their ability to cause zoonotic infections (Butt et al., 2005; Guo et al., 1999; Koopmans et al., 2004).

A number of distinct molecular changes in all gene segments have been shown to play a pivotal role in adaptation, persistence and increased virulence in both poultry and mammalian hosts (Baigent and McCauley, 2003; de Wit et al., 2010; Hatta et al., 2001; Reading et al., 2009; Sorrell et al., 2010; Wu et al., 2009; Yen et al., 2009). Adaptation within a new host species has been notably observed with respect to the two virus glycoproteins which are responsible for virus binding to sialic acid receptors on the host cell surface, fusion of the virus with the cell (both mediated by HA) and release of assembled progeny virus from the infected cell (mediated by NA). During interspecies transfer of LPAI viruses from wild birds to poultry and circulation within a new host species, the HA genes of H5 and H7 viruses can acquire (1) additional glycosylation sites adjacent to the receptor binding site (Banks and Plowright, 2003; Banks et al., 2001; Matrosovich et al., 1999) and (2) multiple basic amino acid residues (arginine and lysine) at the HA cleavage site, which is usually a prerequisite for high pathogenicity. These changes can also be associated with alterations to the NA glycoprotein in which a deletion of several

* Corresponding author. Fax: +44 1635 577263.
E-mail address: munir.iqbal@iah.ac.uk (M. Iqbal).

0042-6822/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2012.08.001

amino acids within the stalk domain is observed (Gulati et al., 2009; Matrosovich et al., 1999). The combination of these adaptive changes in the HA and the NA are thought to modify the complex interactions between the virus receptor (sialic acid), the receptor binding protein (HA) and the NA, the protein responsible for release of the virus from its receptor, therefore altering the balance between the three factors that can result in differences in replication characteristics. These differences alter virus replication in cultured cells (Baigent and McCauley, 2001; Mitnaul et al., 2000; Wagner et al., 2002) and can result in changes in virulence for chickens and mice (Matsuoka et al., 2009; Munier et al., 2010; Reading et al., 2009).

The N-linked carbohydrate side chains at certain positions on the HA of influenza A viruses isolated from various animals and humans are highly conserved and therefore appear to be essential for the formation or maintenance of functional HA (Schulze, 1997). However, oligosaccharide diversity might have a major selective effect and the presence or absence of oligosaccharides on glycosylation sites can determine the fitness of the virus for growth by modulating the biological properties of the HA by interfering with antibody binding (Munk et al., 1992; Schulze, 1997; Skehel et al., 1984), receptor binding (Klenk et al., 2002; Marinina et al., 2003; Wagner et al., 2000), proteolytic activation (Deshpande et al., 1987) and trimer assembly (Roberts et al., 1993).

Among H5 subtype viruses two potential N-linked glycosylation sites at asparagine residues 131 and 158 (H3 numbering) located at the tip of the globular head of HA1 close to the receptor binding site are highly variable (Banks and Plowright, 2003; Matrosovich et al., 1999). An equally striking variability in glycosylation has been documented in H7 subtype viruses (Banks and Plowright, 2003), notably including LPAI H7N1 viruses isolated during circulation in poultry over a nine month period in Italy during 1999–2000. From this epizootic, of 45 HA sequences of LPAI viruses retrieved from public databases, 7 isolates contained a signal for glycosylation at position 123 (133 in H3 numbering) and lacked a motif at position 149 (158 in H3 numbering); the HA from 17 isolates had a glycosylation motif at 149 but not at residue 123, and the remaining 21 viruses had no glycosylation sequon at either site. In contrast, later in the epizootic, following the emergence of HPAI isolates, of 35 HPAI viruses sequences available, the HA of 34 viruses had a glycosylation signal at the 123 site but not at 149 and the remaining isolate lacked a glycosylation motif at either site. The evolutionary drive that led these viruses to acquire additional carbohydrate moieties on the HA and become established in poultry is not clear but modulation of receptor binding might provide these viruses with a greater potential for infection and transmission, as suggested by Matrosovich et al. (1999); alternatively, masking of an antigenic site might make it possible for re-infection of poultry previously infected with viruses of the same sub-type.

In the present study, we sought to investigate the evolution of glycosylation sequons in HA1 of H7N1 viruses in poultry infections and how the virus might change and adapt in an individual poultry host during the course of a single infection. In particular, we asked: (1) are distinct glycosylated variants equally selected and shed from an infected bird and (2) are the differently glycosylated viruses equally well transmitted to contact birds? We established experimental infections of chickens and turkeys with selected H7N1 (LPAI) viruses isolated early during the outbreak in poultry in Italy in 1999 that had conserved glycosylation sites at Asn residues located at positions 12, 28 and 231 in HA1 and residues 82 and 154 in HA2, but the glycosylation at residues 123 and 149 of HA1 was variable. Of the three selected model viruses examined (1) A/turkey/Italy/3466/99 (Italy/3466) encoded a glycosylation site at 123 but not at position 149, (2) A/turkey/Italy/4042/99 (Italy/4042) had glycosylation sites at residue 149 and lacked the site at 123 and (3) A/chicken/Italy/1279/99 (Italy/1279) was found to represent a mixed population with viruses showing four patterns of glycosylation at the two sites with the HA gene coding either no glycosylation at either site, glycosylation at one or the other of the two sites and with glycosylation at each of the two sites. Chickens or turkeys were infected with these viruses and sequence analysis of the HA gene was carried out on multiple cDNA clones prepared from swab samples taken from infected birds to determine the degree of variation or selection that might occur in the virus during infection.

Results

Genetic differences within virus samples used as inocula

To investigate the extent of variation of viruses within infected chicken and turkey hosts during the course of infection both species were infected experimentally with three H7N1 (LPAI) viruses, Italy/3466, Italy/4042 and Italy/1279. These viruses had been chosen on the basis of variation at the N-linked potential glycosylation sequons, as described above, in the vicinity of the receptor binding site of the HA. The reported sequences for Italy/3466 had a potential glycosylation site at residue Asn 123 but lacked a site at residue Asn 149 (mature H7 HA1 numbering, open reading frame numbers Asn 141 and Asn 167 respectively). Italy/4042 contained a potential glycosylation motif at residue 149 but not at position 123. The reported sequence for Italy/1279 virus lacked glycosylation motifs at both positions. Sequences of the remaining gene segments of these viruses were not available in public databases.

To analyse the genetic difference between these viruses, sequence analysis of all eight gene segments per virus was undertaken. cDNA amplicons covering the complete open reading frames of PB2, PB1, PA, HA, NP, NA, M and NS genes were prepared from allantoic fluid made for inoculum stocks; cDNAs were cloned and between 6 and 10 individual cDNA clones prepared from each segment for each virus were sequenced. Sequence analysis based on this number of cDNA clones analysed showed very limited polymorphism within each gene segment for each virus (Table 1) with the exception of the HA gene of Italy/1279 virus which showed marked polymorphism. This heterogeneity was observed within the inoculum at nucleotide positions 427 and 505 resulting in amino acid changes (A→T or vice versa) at residues 125 and 151 which alters the potential glycosylation sequon at amino acid residues Asn 123 and Asn 149. The level of heterogeneity of the HA1 coding region of each virus was examined in greater detail using a ‘deep amplicon’ sequencing approach.

Identification of glycosylation status at residues Asn 123 and Asn 149

In order to determine whether potential glycosylation motifs at position 123 and 149 were able to be glycosylated, four recombinant reverse genetics (RG) viruses containing HA and NA gene of Italy/1279 virus were generated. The HA gene of RG viruses encoded the four different potential glycosylation patterns observed in the Italy/1279 inoculum stock: (1) glycosylation at residue 149 and not at residue 123, (2) glycosylation at residue 123 and not at residue 149, (3) glycosylation at both residues 123 and 149 and (4) no glycosylation at either site. SDS-PAGE analysis of HA1 of these RG viruses showed differential mobility depending on the potential glycosylation status (Fig. 1A). The electrophoretic mobility of the HA1 polypeptide of viruses with potential glycosylation at residue 149 (Fig. 1A, lane 1) or at residue 123

Table 1
Differences in amino acid residues between Italy/3466, Italy/4242 and Italy/1279 viruses used as inocula to infect chickens and turkeys.

Gene product	No. of variable amino acids	Differing residues at position	Amino acid substitutions		
			Italy/ 3466	Italy/ 4042	Italy/ 1279
PB2	3	87	**G**	D	D
		215	T	T	**I**
		508	R	R	**M**
PB1	3	74	P	**S**	P
		433	K	**R**	K
		674	I	**V**	I
PB1-F2	1	69	Q	**R**	Q
PA	3	269	R	**K**	R
		339	**G**	E	E
		401	R	R	**I**
HA[1]	10	−5	I	**T**	I
		60	I	**T**	I
		125	T	A	**A/T**
		151	**A**	**T**	**A/T**
		155	M	**L**	M
		177	G	**V**	G
		188	S	**G**	S
		201	**L**	Q	Q
		207	S	**I**	S
		208	P	**L**	P
NP	0	–	–	–	–
NA	4	43	Q	**R**	Q
		95	I	**M**	I
		179	G	**E**	G
		266	I	**V**	I
M1	1	107	**V**	I	I
M2	1	16	E	**V**	E
NS1	5	45	G	G	**R**
		91	T	**I**	T
		204	**E**	G	G
		221	Y	**H**	Y
		228	S	**P**	S
NS2	2	12	I	I	**T**
		47	**K**	E	E

Amino acids variants between Italy/1279, Italy/3466 and Italy/4042 viruses are indicated in bold. [1]HA numbering is of the mature H7 HA glycoprotein.

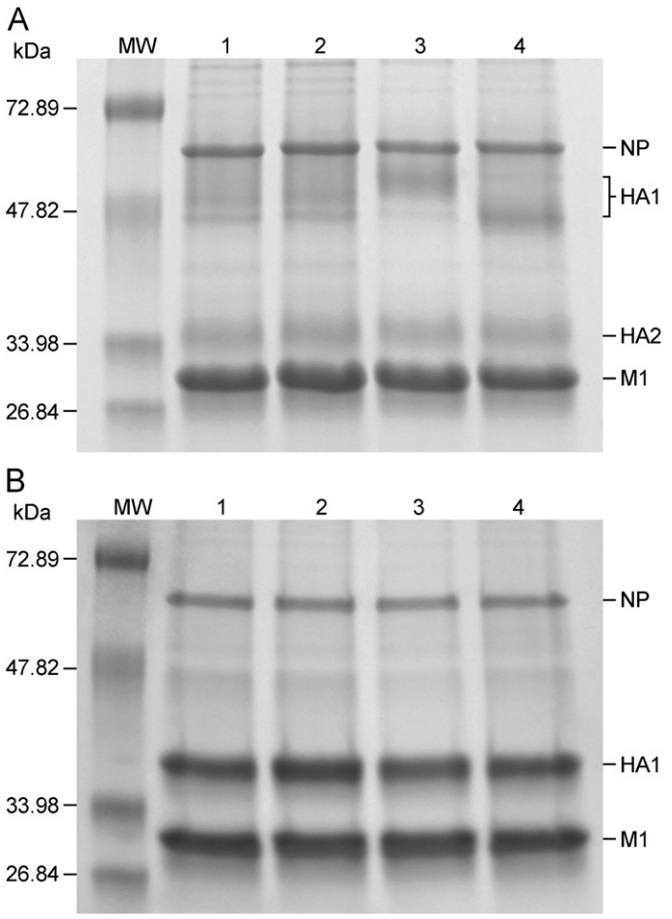

Fig. 1. Analysis of the electrophoretic mobility of the HA1 of RG viruses that (A) differ in their potential glycosylation sites at residues Asn 123 and Asn 149 in the HA of Italy/1279 and (B) treated with PNGase F. Purified reassortant RG containing HA and NA of Italy/1279 and internal genes of PR8/34 (H1N1) were analysed on SDS-PAGE under reducing conditions. Protein bands were stained with Coomassie blue. HA1 of RG viruses contained: (lane 1), a glycosylation motif at residue 149 but not at residue 123; (lane 2), a glycosylation motif at residue 123 but not at residue 149 site; (lane 3), glycosylation motifs at both residues 123 and 149, and (lane 4), no glycosylation motifs at either residue 123 or residue 149.

(Fig. 1A, lane 2) was intermediate compared with the virus that contained glycosylation at both residues 123 and 149 (Fig. 1A, lane 3), which showed a reduced electrophoretic mobility, and with the virus that contained no potential glycosylation motifs at either of these two sites (Fig. 1A, lane 4), the HA1 of this virus showing the fastest mobility. Importantly, HA1 of each virus generated by reverse genetics showed equal electrophoretic mobility to each other following deglycosylation by PNGase F treatment (Fig. 1B). These results indicated that viruses with potential glycosylation signals at both residues 123 and 149 of the HA are able to be glycosylated at both of these sites in avian cells.

Identification of virus variants with differing HA1 sequences in the virus stocks used as inocula

With an optimised RT-PCR procedure, sequence analysis of 200 cDNA clones of the HA1 coding region of the HA gene in Italy/3466 virus inoculum showed that 78% of the cDNA clones had an identical nucleotide sequence to the consensus sequence of the inoculum, with nucleotide variation at a frequency of over 1% seen at only two sites (Fig. S1A). This high degree of sequence conservation was not seen in cDNA clones of the HA1 coding regions produced from Italy/4042 and Italy/1279 which showed an increased number of synonymous and non-synonymous nucleotide substitutions compared with their derived consensus sequences. Analysis of 93 of the cDNA clones of the HA1 coding region derived from the Italy/4042 inoculum showed that 35% of clones shared their nucleotide sequences with that of the consensus sequences of the inoculum virus with 42% of clones differing by a single nucleotide from the consensus; two sites showed over 5% variation within the set of cDNA clones when compared with the consensus sequence of the inoculum (Fig. S2A). For the Italy/1279 inoculum, a set of 126 clones produced corresponding to the HA1 coding region showed marked polymorphism at nucleotides 427 and 505 of the HA gene, observed at an equal frequency of 37.4%. Of the clones analysed, 13% were identical with the consensus sequence of the inoculum and 63% differed by a single nucleotide from the inoculum consensus sequence; three individual sites showed heterogeneity within the set at greater than 5% when compared with the consensus sequence of the inoculum, nucleotides 427, 505 and 680 (Fig. S3A). Table 2 lists the amino acid substitutions resulting from non-synonymous nucleotide substitutions observed at a level of 5% or over compared with the derived consensus sequence of the cDNA clones produced from RNA extracted from the inoculum.

Table 2
Encoded variable amino acid residues in HA1 during infection of chickens and turkeys inoculated with Italy/3466, Italy/4042 and Italy/1279 observed at greater than 5% heterogeneity.

Virus strain	Host		dpi	Swabs/No. of birds	No. of clones	Inoculum consensus sequence[1].									
						A125[a]	A128	S132	A151[a]	L192	S207	G209	N231	S249	V314
						Number of amino acid substitutions and polymorphisms (%)									
Italy/3466	Inoculum				200										
	Chickens		4	Buccal/10											
				Cloacal/2	788										
Italy/4042	Inoculum				93						15I (15.6)				11I (11.9)
	Chickens		6	Buccal/10											
				Cloacal/8	303						19I (6.3)				265I (87.4)
Italy/1279	Inoculum				126	**47T (37.4)**			**47T (37.4)**			10E (7.9)			
	Chicken (infected)		1	Buccal/7	148	**108 T (72.9)**		15L (10.1)	**38T (25.7)**				15S (10.1)		
			4	Buccal/8	74	**71T (95.9)**	16T (21.6)								
			5	Cloacal/3	48	**34T (70.8)**			**14T (29.2)**						
			12	Cloacal/1	15	**15T (100)**		15L (100)					15S (100)		
			13	Buccal/1	19	**19T (100)**									
				Total	304	**246T (80.9)**	16T (5.3)	30L (9.9)	**54T (17.8)**				30S (9.9)		
	Turkeys (infected)														
			1	Buccal/8	185	**78T (42.2)**			**104T (56.2)**						
			4	Buccal/7											
				Cloacal/3	106	**85T (80.2)**			**9D (8.5)**						
			6	Buccal/5											
				Cloacal/1	82	**79T (96.4)**									
				Total	373	**242T (64.9)**			**104T (27.9)**						
	Turkeys (infected and contact group)	Infected	3	Buccal/ 8											
				Cloacal/4	921	**546T (59.3)**			**314T (34.1)**						
									51D (5.5)						
		Contact	3	8/Buccal	589	**584T (99.2)**									
						7T (1.2)								358N (60.8)	
			9	8/Buccal											
				4/Cloacal	995	**984T (97.9)**									
										103M (10.4)					
														516N (51.8)	
				Total	1584	**158T (99.8)**				106M (6.7)				826N (52.2)	

[a] Variation at amino acids 125 and 151 (indicated in bold) alter the potential glycosylation signal at Asn residues 123 and 149 respectively. The number in the parentheses is the percentage of variants that differ from the consensus sequences of their corresponding inoculum sequence. [1]HA numbering is of the mature H7 HA glycoprotein.

Sequences that show less than 1% synonymous or non-synonymous substitutions compared with the derived consensus sequence from the inoculum clones have not been shown in the figures.

The results of the deep amplicon sequence analysis confirmed that there was a pronounced polymorphism at nucleotides 427 and 505 of the HA gene of Italy/1279. The nucleotide polymorphism indicated that Italy/1279 virus contained a combination of four glycosylation signals at two critical Asn residues (123 and 149) in the HA: (1) 27% of clones did not encode a signal for glycosylation at either of the two key residues of the HA, (2) 71.4% of clones encoded a single glycosylation motif, encoding glycosylation at either residue 123 (35.7% of clones) or 149 (35.7% of clones), and (3) 1.6% of the total clones encoded a glycosylation motif at both sites of the HA. The other two viruses studied, Italy/3466 and Italy/4042, each encoded a homogeneous glycosylation motif at these two sites. All 200 cDNAs corresponding to the HA1 coding region of the HA gene sequenced from the Italy/3466 inoculum encoded a glycosylation motif at position 123 but did not encode a motif at residue 149 whereas all 93 cDNAs corresponding to the HA1 coding region of the HA gene derived from Italy/4042 virus encoded motifs for glycosylation at Asn 149 but did not encode a motif for glycosylation at Asn residue 123 (Table 3).

Table 3
Analysis of the encoded variable glycosylation motifs at residues 123 and 149 in the HA1 globular head domain of H7N1 viruses used as inoculum.

Inoculum virus strain	Total no of cDNA clones sequenced	Number of clones (%)	Glycosylation at residue	
			123[1] (133)[2]	149[1] (158)[2]
Italy/3466	200	200 (100)	+	−
Italy/4042	93	93 (100)	−	+
Italy/1279	126	45 (35.7)	+	−
		45 (35.7)	−	+
		34 (27.0)	−	−
		2 (1.6)	+	+

[1]HA numbering is of the mature H7 HA glycoprotein. [2]H3 numbering of the H3 HA glycoprotein. (+) indicates the presence of a glycosylation motif and (−) represents the absence of a glycosylation motif.

Analysis of virus variants in infected chickens

Differentially glycosylated viruses (Italy/3466, Italy/4042 and Italy/1279) were used separately to infect chickens; groups of 10 animals were infected experimentally with one of three different doses, nominally 10^2, 10^4 and 10^6 50% egg infectious dose (EID_{50}) per virus. The animals were sampled daily with swabs taken from the buccal and cloacal cavities and the presence and duration of virus shedding was determined by quantitative real-time PCR. In the group of birds infected with the lowest dose (10^2 EID_{50}) less than 100% of birds became infected and the group showed variable infection profiles (Table S1). No marked differences in the ability of three viruses to infect the birds were seen between the three virus groups when the birds were infected with a 10^4 or 10^6 EID_{50} virus dose (Tables 4–6).

Sequence heterogeneity of the HA1 coding region of the HA gene of virus present in the buccal and cloacal swabs taken from chickens infected with Italy/3466 (Tables 4A, 4B), Italy/4042 (Table 5B) and Italy/1279 (Table 6B) at doses of 10^4 or 10^6 EID_{50} was examined in samples with adequate levels of virus RNA as determined by quantitative RT-PCR; no analysis was done on swabs taken from birds infected with the lowest dose due to the low levels of RNA detected in the swabs taken from these birds. Production of PCR products from swab samples with below 10^2 relative equivalent units (EID_{50}) per ml were not reliable using 12% of the RNA taken from the standard 140 μl sample volume of a swab. With this limit of detection, swabs chosen for analysis had RNA levels equivalent to $\geq 10^2$ EID_{50}/ml. In light of the possibility that greater polymorphism or selection of variants within individual infected birds might be seen in virus samples recovered towards the end of the infection, birds were sampled on days, as late as possible in infection, on which the majority of birds shed virus at an adequate level for deep amplicon analysis. This resulted in swabs being chosen for analysis from birds as early after infection (24 h) for all viruses, and as late as feasible: from the group infected with Italy/3466 at 4 days post-infection (dpi) (Tables 4A, 4B) and the group infected with Italy/4042 at 6 dpi (Table 5B). For infections with Italy/1279, the virus was known to have polymorphism in the inoculum stock thus sampling was done during the middle of the infectious period

Table 4A
Detection of RNA titres in buccal and cloacal swabs collected over 21 days from chickens inoculated with Italy/3466 virus at $10^{3.8}$ EID_{50}.

Chicken ID	Swab	Days post infection																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
64	Buccal	2.2	2.4	2.0	–	3.3	2.5	–	–	2.2	–	–	–	–	–	2.4	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	–	1.6	–	–	–	–	–	–	–	–	–
65	Buccal	2.2	3.2	2.5	**3.3**	1.7	2.9	2.2	2.2	–	–	2.5	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	1.6	**3.0**	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
66	Buccal	3.2	3.0	2.2	**3.4**	2.3	3.3	2.5	2.1	2.1	–	1.5	1.5	–	–	–	–	–	1.5	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	2.4	–	–	–	–	–	–	–	–	–	–
67	Buccal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
68	Buccal	2.0	2.8	2.2	**3.4**	2.4	1.6	2.1	–	–	–	2.2	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	2.1	1.6	**3.3**	2.3	3.1	2.4	2.3	3.5	2.2	2.1	2.5	–	1.6	2.4	1.6	2.3	2.1	–	–	2.1
69	Buccal	2.4	2.6	2.4	**3.6**	3.3	2.3	2.1	1.6	–	1.6	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	1.5	2.9	2.1	2.2	2.2	2.4	1.6	1.6	–	–	–	1.6	3.1	2.1	–	–	–	–
70	Buccal	2.9	2.0	2.4	**3.8**	2.5	2.6	1.6	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	1.5	2.5	2.5	3.0	3.2	2.7	1.7	1.5	2.9	–	–	1.6	1.6	–	–	–	–	–
71	Buccal	–	–	–	–	–	–	–	–	–	–	2.5	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
72	Buccal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
73	Buccal	3.4	2.1	2.9	**3.9**	3.1	2.6	1.8	1.6	1.6	–	2.9	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	1.6	2.0	–	1.5	–	–	–	2.3	–	–	–	–	–	–	–	–	–	–

Viral RNA levels in swabs were detected by real-time RT-PCR. Measured Ct values were extrapolated as relative equivalent units (REU) of virus infectivity titres in $\log_{10}$ EID_{50} per ml based on a standard curve constructed from RNA extracted from known EID_{50} titres of Italy/3466. RNA titres indicated in bold were selected for deep amplicon sequence analysis. (−) indicates RNA titres $< 10^1$ EID_{50} per ml.

Table 4B
Detection of RNA titres in buccal and cloacal swabs collected over 21 days from chickens inoculated with Italy/3466 virus at $10^{6.2}$ EID_{50}.

Chicken ID	Swab	Days post infection																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
EE2C9	Buccal	3.3	3.8	2.5	**4.1**	2.1	2.9	1.5	–	–	–	1.6	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	2.4	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
EFC70	Buccal	3.2	2.8	2.6	**3.7**	2.3	2.0	2.1	–	–	–	–	–	2.1	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–		3.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–
EF3EF	Buccal	4.2	2.8	2.6	**3.3**	2.4	2.5	2.0	–	–	–	2.3	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	2.3	–	4.0	3.3	3.7	–	2.7	2.9	2.7	2.5	2.7	2.3	–	2.5	2.0	2.0	–	–	–	–
EEF34	Buccal	3.0	2.9	2.7	**3.7**	2.4	2.3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	1.5	3.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–
F16CC	Buccal	3.6	2.7	2.7	3.7	2.3	2.2	2.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	3.1	–	2.6	2.2	3.0	2.8	3.2	2.8	2.7	2.4	–	–	–	–	–	–
EDFDB	Buccal	3.2	2.5	2.4	3.4	2.3	2.5	2.7	1.6	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	2.7	–	–	–	–	–	–	–	–	–	–	–	–	–	–
EDAE1	Buccal	3.7	2.6	2.7	3.8	2.5	2.5	–	2.2	2.0	2.0	–	2.1	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	2.3	3.7	2.7	–	–	3.2	2.7	2.5	2.5	3.4	2.6	–	2.8	2.5	–	–	3.0	2.4	–
F030C	Buccal	3.9	2.5	2.2	3.8	2.3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
EE0D2	Buccal	3.2	3.1	3.6	3.3	2.4	–	2.1	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	2.2	3.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–
EE152	Buccal	3.8	2.7	2.6	3.5	2.2	3.2	–	–	–	1.5	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	2.4	2.2	2.2	3.5	2.3	3.2	2.1	3.5	2.5	2.8	2.0	2.6	2.2	–	–	–	–	–	–	–	–

RNA titres indicated in bold were selected for deep amplicon sequence analysis. (–) indicates RNA titres $< 10^1$ EID_{50} per ml. Viral RNA levels in swabs were detected by real-time RT-PCR as described for Table 4A.

Table 5A
Detection of RNA titres in buccal and cloacal swabs collected over 21 days from chickens inoculated with Italy/4042 virus at 10^4 EID_{50}.

Chicken ID	Swab	Days post infection																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1341	Buccal	2.0	2.5	2.4	2.3	2.8	2.7	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1.6	–	–	–	–	–
1342	Buccal	2.6	3.3	4.0	3.1	3.1	3.0	2.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	2.4	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1343	Buccal	4.4	3.2	4.5	3.6	3.3	2.8	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1344	Buccal	3.2	4.1	3.4	3.2	3.1	2.6	1.9	2.2	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	1.6	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1345	Buccal	3.0	3.2	3.2	3.1	3.0	2.2	2.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	5.9	2.5	2.9	2.1	–	–	–	–	–	–	–	–	–	–	–	–	–
1346	Buccal	3.3	2.9	3.1	3.0	3.9	3.1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	1.6	–	1.6	2.9	3.8	3.1	–	2.8	–	–	–	–	–	–	–	–	–	–	–	–	–
1347	Buccal	3.6	2.9	5.0	4.3	3.2	2.7	2.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1348	Buccal	2.2	2.9	3.8	3.5	3.5	2.7	2.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	1.6	2.5	3.6	–	2.7	–	–	–	–	–	–	–	–	–	–	–	–	–
1349	Buccal	2.9	3.0	3.3	2.9	3.0	2.5	–	–	–	–	–	–	–	–	–	1.6	–	–	–	–	–
	Cloacal	–	–	–	–	–	1.6	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1350	Buccal	3.3	3.4	5.1	2.9	3.3	3.0	1.6	2.1	–	–	–	2.2	–	–	–	–	–	–	–	–	–
	Cloacal	–	5.4	2.2	2.7	2.7	3.0	2.8	–	–	–	–	–	–	–	–	1.6	–	–	–	–	-

Viral RNA levels in swabs were detected by real-time RT-PCR as described for Table 4A but using RNA from known titres of Italy/4042 as the standard. (–) indicates RNA titres $< 10^1$ EID_{50} per ml.

(4–5 dpi) (Table 6B). RNA was prepared from swabs taken from individual birds at various time points and for each RNA multiple cDNA clones corresponding to the HA1 coding region of the HA gene were prepared and sequenced.

Changes in the HA1 during infection of chickens with Italy/3466 and Italy/4042

A total of 10 buccal and 2 cloacal samples taken on 4 dpi from 10 birds infected with Italy/3466 (Tables 4A, 4B), were used to prepare a total of 788 cDNA clones; these showed only very limited heterogeneity in the HA1 coding region. Similar to the heterogeneity observed within the inoculum, 75% of the cDNA clones were identical to the consensus sequence of the inoculum virus and 17% of the cDNA clones showed only a single nucleotide polymorphism compared with the consensus sequence of the virus used for inoculation. A single site showed heterogeneity with 2.6% of the total clones, but all variants observed were from a single bird and represented a frequency of 87% of the cDNAs analysed from this bird (Table S2A). This site encoded a

Table 5B
Detection of RNA titres in buccal and cloacal swabs collected over 21 days from chickens inoculated with Italy/4042 virus at $10^{5.4}$ EID_{50}.

Chicken ID	Swab	Days post infection																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1351	Buccal	2.2	1.9	3.8	3.8	2.8	**2.9**	–	–	–	–	–	–	–	–	–	–	1.6	–	–	–	–
	Cloacal	–	–	2.8	3.5	3.5	**3.8**	–	2.4	–	–	2.1	–	3.2	2.9	–	–	2.2	–	–	2.5	1.6
1352	Buccal	1.8	3.0	3.1	3.1	2.3	**2.9**	2.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	3.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1353	Buccal	2.7	3.2	3.7	3.8	3.6	**2.8**	–	1.6	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	4.4	3.6	**3.3**	2.3	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1354	Buccal	–	3.4	4.3	3.4	2.9	**2.6**	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	2.1	–	–	–	4.7	**3.5**	3.8	2.1	–	–	–	–	–	–	–	–	–	–	–	–	–
1355	Buccal	1.8	3.0	2.6	3.1	1.9	**2.8**	2.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	3.4	**3.5**	–	2.3	–	–	–	–	–	–	–	2.4	2.2	2.7	–	–	–
1356	Buccal	2.3	3.3	3.6	4.5	3.1	**2.7**	2.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	1.9	–	3.1	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1357	Buccal	2.7	2.8	3.5	4.2	3.3	**3.2**	2.3	2.2	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	4.1	**3.5**	3.1	–	3.1	–	–	–	–	–	–	–	–	–	–	–	–
1358	Buccal	2.0	2.5	4.1	2.5	1.7	**3.3**	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	2.0	–	1.3	**3.3**	3.8	–	–	–	–	–	–	–	–	2.1	–	–	–	–	–
1359	Buccal	2.6	3.3	3.7	3.3	1.6	**2.7**	2.4	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	3.7	**4.2**	2.1	2.4	–	–	–	–	–	–	–	–	1.6	–	–	–	–
1360	Buccal	1.6	2.7	3.4	3.3	1.5	**2.5**	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	3.7	4.7	**2.9**	–	1.6	–	–	–	–	–	–	–	–	2.3	–	–	–	–

RNA titres indicated in bold were selected for deep amplicon sequence analysis. (-) indicates RNA titres $< 10^1$ EID_{50} per ml. Viral RNA levels in swabs were detected by real-time RT-PCR as described for Table 5A.

Table 6A
Detection of RNA titres in buccal and cloacal swabs collected over 21 days from chickens inoculated with Italy/1279 virus at $10^{3.8}$ EID_{50}.

Chicken ID	Swab	Days post infection																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1542	Buccal	2.5	2.5	2.6	3.0	–	3.3	2.9	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	2.1	–	–	2.2	2.3	–	–	2.5	2.2	–	–	–	–	–	–	–	–	–
E6FCB	Buccal	2.3	2.5	–	2.6	2.2	2.8	2.8	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	3.6	3.0	3.5	2.3	2.6	2.4	2.3	2.1	–	–	–	–	–	–	–	–	–
28A91	Buccal	5.2	2.8	–	2.5	2.2	2.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	2.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
85634	Buccal	2.1	2.6	2.3	3.2	3.7	2.6	3.6	2.4	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	2.8	–	–	2.0	-	2.2	–	–	–	–	–	–	–	–	–
1541	Buccal	3.3	2.7	2.0	2.9	2.7	2.5	2.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	2.0	–	2.2	2.1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
867D2	Buccal	3.2	2.6	3.2	3.0	2.6	2.5	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	3.3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
9AF2E	Buccal	–	2.1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
F840D	Buccal	2.0	2.7	2.3	2.9	2.5	3.1	2.6	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
F0808	Buccal	2.9	2.6	2.0	2.6	2.3	3.3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	2.2	3.4	2.3	2.3	2.4	2.4	–	–	–	–	–	–	–	–	–	–
994F9	Buccal	3.0	–	–	2.6	2.3	2.8	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	2.3	3.1	4.5	4.0	3.0	2.2	2.0	–	–	–	–	–	–	–	–	–	–	–

Viral RNA levels in swabs were detected by real-time RT-PCR as described for Table 4A but using RNA from known titres of Italy/1279 as the standard. (–) indicates RNA titres $< 10^1$ EID_{50} per ml.

polymorphism encoding alanine and threonine at residue 97 of the HA although no variation at this site was observed in the inoculum virus (Fig. S1). These results showed that very few variants were detected during infection of chickens with Italy/3466.

A total of 303 cDNA clones corresponding to the HA1 coding region of the HA gene derived from 10 buccal and 8 cloacal samples taken on 6 dpi from 10 chickens infected with Italy/4042 were analysed (Table 5B). In contrast to the results obtained from chickens infected with Italy/3466, only 4% of cDNA clones were identical to the consensus sequence of the cDNAs prepared from the inoculum; 92% showed variation at a single nucleotide. This variation resulted in two non-synonymous nucleotide substitutions, encoding S207I and V314I detected at a frequency of 6.3% and 87.5% respectively in the cDNA clones; sequence heterogeneity had been observed to encode changes at these sites in the HA sequence determined from the inoculum, at a frequency of 15.6%, and 11.9% respectively encoding amino acid substitutions S207I/N, V314I (Table 2). Variation at residue 207 was seen in three birds (at a range of 3.5% and 48% of the cDNAs analysed,

Table 6B
Detection of RNA titres in buccal and cloacal swabs collected over 14 days from chickens inoculated with Italy/1279 virus at $10^{6.6}$ EID_{50}.

Chicken ID	Swab	Days post infection												
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Buccal	**3.9**	2.6	3.1	**3.7**	–		2.1	–	–		–	–	2.7	
	Cloacal	–	–	–	–	2.2		–	–	–		–	–	–	
2	Buccal	**3.3**	1.5	2.3	**2.7**	–		1.5	–	–		–	2.4	2.4	
	Cloacal	–	–	–	–	**2.7**		3.1	-	3.3		2.7	2.6	–	
3	Buccal	2.7	2.2	2.2	2.4	–		2.2	–	2.2		–	–	–	
	Cloacal	–	–	–	–	–		–	–	–		–	–	–	
4	Buccal	3.3	2.1	2.1	2.4	-		1.2	–	–		–	2.8	–	
	Cloacal	–	–	–	–	–		–	–	-		–	–	–	
5	Buccal	**4.0**	2.5	2.3	**3.2**	–		1.5		–		–	2.3	–	
	Cloacal	–	–	–	–	–		–	1.6	–		–	–	–	
6	Buccal	**3.6**	2.4	2.3	**3.3**	–		2.1	–	–		–	2.7	–	
	Cloacal	–	–	–	–	**3.4**		3.0	–	2.4		–	–	2.4	
7	Buccal	**3.9**	2.5	2.1	**2.8**	2.5		2.4	–	–		–	2.4	2.4	
	Cloacal	–	–	3.2	4.8	**3.5**		2.9	2.8	2.8		2.2	2.9	2.7	
8	Buccal	**3.8**	2.4	2.5	**3.3**	2.6		2.1	1.3	–		–	–	2.8	
	Cloacal	–	–	–	–	–		–	–	–		–	–	–	
9	Buccal	2.3	2.2	3.9	**3.1**	–		1.5	–	2.6		–	–	2.8	
	Cloacal	–	–	–	–	–		1.5	–	–		–	–	–	
10	Buccal	**4.2**	2.1	3.3	**3.2**	2.9		1.2	–	2.5		–	–	2.5	
	Cloacal	–	–	–	–	2.4		3.1	2.4	2.5		–	–	–	

Viral RNA levels in swabs were detected by real-time RT-PCR as described for Table 6A. RNA titres indicated in bold were selected for deep amplicon sequence analysis. Empty squares indicate that the swabs were not analysed. (–) indicates RNA titres $< 10^1$ EID_{50} per ml.

Table S2A) but samples prepared from five of the ten birds showed variation at residue 314 with a range in individual birds showing polymorphism of between 48% and 97% (Table S2A). Three additional minor (1–2%) substitutions were also observed in cDNAs prepared from the swab samples (Fig. S2), two of which, nucleotides 543 and 836, were also present in the inoculum. The polymorphism at 543 was a synonymous substitution. The changes at nucleotides 836 and 906 encoded amino acid substitutions, with a resulting change G261E as a minor population in a single bird (Table S2A) and P284S present as a minor population (1.7–7.6% of cDNAs) in three birds. The observations that there was an increase in the substitution V314I and a reduction in the substitution S207I in clones prepared from swabs taken from infected birds suggested that positive selection had occurred at position 314 and negative selection at position 207 during replication in chickens but the significance of these substitutions is not known .

Changes in the HA1 during infection of chickens with Italy/1279

Much more striking genetic variation was observed in cDNA clones corresponding to the HA1 coding region of the HA gene derived from swab samples from birds inoculated with Italy/1279. This virus had marked heterogeneity at nucleotide positions 427 and 505 altering two glycosylation sites in the vicinity of the receptor binding site of the HA. To examine possible selection of these variants in the infected host samples were taken from 7 birds on dpi 1, from 8 birds on dpi 4, from 3 birds on dpi 5 (Table 6B), and analyses of 270 cDNA clones representing the course of infection revealed a large number of non-synonymous substitutions. Within any single time point nucleotide positions 427, 436, 505, 697, 718 and 866 showed $>1\%$ heterogeneity in comparison with the inoculum virus (Italy/1279) consensus sequence (Fig. S3) and encode 125A/T, 128T/A, 151T/A/D, 215N/D, 222D/N, and 271D/G in the HA (Table S2B, Table 2). The polymorphic nucleotide changes encoding variation at amino acid residues 125 and 151 affect the glycosylation sites at Asn 123 and Asn 149, respectively, and represented the major sites of sequence heterogeneity between the clones.

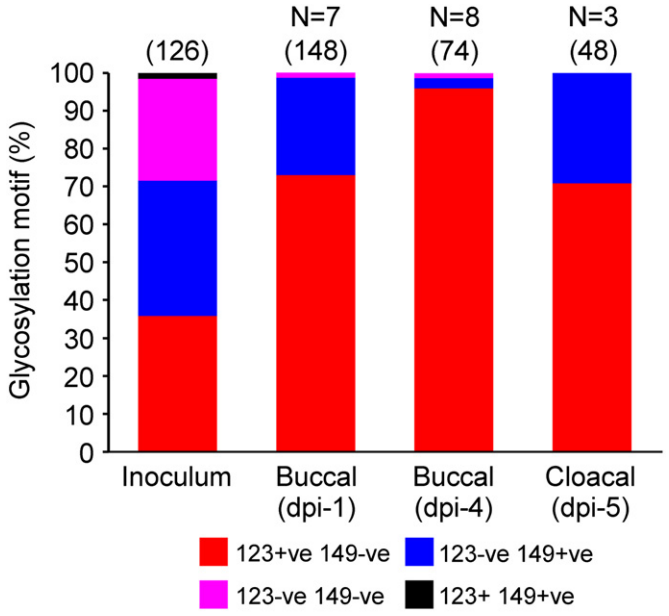

Fig. 2. Changes in glycosylation motifs in the HA1 polypeptide during infection of chickens with $10^{6.6}$ EID_{50} Italy/1279. The presence (+ve) and absence (-ve) of glycosylation sites at residues 123 and 149 in HA1 are shown as 123+ve and 149-ve (red bars), 123-ve and 149+ve (blue bars), 123-ve and 149-ve (pink bars) and 123+ve and 149+ve (black bars). The bars represent the percentage of sequences encoding a specific pattern of glycosylation at Asn 123 and 149 sites at the indicated time points. *N* indicates the number of birds sampled at each time point and the number of clones analysed at each time point are indicated in parenthesis.

Changes in the glycosylation pattern of the HA over the course of infection of chickens with Italy/1279

Sequence data derived from 148 cloned cDNAs corresponding to the HA1 coding region of the HA gene derived from buccal swab samples recovered from 7 birds at 1 dpi retained a mixed population (Fig. S4A). The vast majority of clones (98.7%) encoded either a glycosylation site at residue 123 or at residue 149, with 75% of clones encoding a glycosylation sequon at residue 123 but not at 149 (Fig. 2). Compared with the inoculum there was a

substantial decrease (from 26.9% to 1.3%) in sequences encoding no glycosylation on either of these two (123 or 149) sites. A decrease in the frequency (from 1.5% to 0%) of sequences encoding a glycosylation site at both 123 and 149 sites was seen but this loss might represent insufficient sampling to detect this minor population. Sequences analysed on 4 dpi (74 clones from buccal samples from 8 birds) and 5 dpi (48 clones from cloacal samples from 3 birds) showed that the majority of the population of viruses encoded a glycosylation site at residue 123. However, one bird on 4 dpi retained a mixed population with approximately 20% viruses encoding glycosylation at 149 site and 10% viruses with no glycosylation on either the 123 or 149 site (Fig. S4B). cDNA amplification on day five was focussed on cloacal samples but was successful for only three birds. Each showed a single population of cDNAs, all cDNAs from one bird encoded glycosylation at HA1 residue 149 but not at 123 of HA1; cDNAs prepared from the cloacal samples of the other two birds directed glycosylation at residue 123 but not 149 of HA1 (Fig. S4C).

Overall the results from the infected chickens indicate that viruses with a single glycosylation motif in the vicinity of the receptor binding site of the HA replicated preferentially in these birds infected with Italy/1279, Italy/3466 and Italy/4042. For Italy/3466 and Italy/4042, with a single glycosylation site in the vicinity of the receptor binding site, there was no evidence for variation at the glycosylation sites. However, when the inoculating virus stock contained several alternative glycosylation patterns in this region of the HA, viruses that encoded glycosylation at residue 123 were found to predominate over viruses encoding glycosylation at residue 149 of the HA and each of the single site glycosylations vastly predominated both over viruses that encoded glycosylation sites at both residues (123 and 149) and over viruses that encoded neither glycosylation motif.

Selection of virus variants in turkeys

The original virus Italy/1279 was isolated from an infected broiler chicken; however, the 1999/2000 epizootic outbreak was predominant in turkeys. We therefore examined whether a different poultry host might show selection of viruses with distinct glycosylation patterns in the HA. For these experiments three groups of turkeys (10 in each group) were infected with Italy/1279 virus at nominal doses of 10^2, 10^4 and 10^6 EID_{50}. Samples (buccal and cloacal swabs) taken on 1, 4 and 6 dpi (Table 7) were subjected to deep amplicon sequence analysis and a total of 373 cDNA clones corresponding to the HA1 coding region of the HA gene were analysed. Non-synonymous variation above the level of 1% of the cDNA clones analysed was detected at nucleotide positions 422, 427, 436, 505, 506, 680 and 1024 of the HA gene (Fig. S5). These nucleotide variants encode the amino acid substitutions 123N/S, 125A/T, 128A/T, and 151A/T/D, 209G/E, and 324F/L in the HA respectively (Table S2C). When examined as heterogeneity greater than 5% over the six days of infection the number of polymorphic sites was reduced and polymorphism only at amino acid residues 125 and 151 (Table 2), altering the glycosylation signal at residues Asn 123 and Asn 149 of the HA as described above, was observed.

Analysis of glycosylation sites in the HA of viruses recovered from buccal and cloacal swabs taken from turkeys infected with Italy/1279 virus showed a similar pattern of glycosylation sites in the HA to that seen during infection of chickens. At 1 dpi (185 clones corresponding to the HA1 coding region of the HA gene from buccal samples from 8 birds), the patterns of glycosylation motifs of the HA were similar to the inoculum, with 7 of the 8 bird analysed showing evidence of a mixed population of virus (Fig. S6A); swabs from the single bird with a homogenous pattern encoded no glycosylation sequon at residue 123 but a glycosylation sequon at 149. Swabs from the other birds had no discernible pattern. By 4 dpi (106 clones from 7 birds, 7 buccal and 3 cloacal samples) swabs from five of seven birds from which cDNAs were successfully produced showed a mixed population of sequences (Fig. S6B). At this time point the pattern of heterogeneity within the birds was less heterogeneous than the pattern seen at day 1 post infection but the pattern was somewhat more variable than the pattern seen in infected chickens at the same time point. Swabs taken from all but one bird showed a predominance of cDNAs encoding glycosylation at residue 123 but not at 149. However, cDNAs prepared from the swabs from one bird (bird

Table 7
Detection of RNA titres in buccal and cloacal swabs collected over 6 days from turkeys inoculated with Italy/1279 virus at $10^{7.4}$ EID_{50}.

Turkey ID	Swab	Days post infection						
		1	2	3	4	5	6	7
31	Buccal	**3.2**	4.2	4.0	4.1	2.9	Died	
	Cloacal	–	–	–	2.3	1.5		
32	Buccal	**3.6**	4.5	3.6	**3.3**	3.6	**2.7**	Killed
	Cloacal	–	–	–	–	–	–	
33	Buccal	**5.3**	5.0	5.1	**4.7**	3.3	**4.2**	Killed
	Cloacal	–	–	2.3	–	–	1.5	
34	Buccal	**4.2**	3.8	4.5	**4.3**	4.0	**3.1**	Killed
	Cloacal	–	–	–	–	1.5	–	
35	Buccal	**4.0**	3.6	4.3	**2.9**	3.7	**3.1**	Killed
	Cloacal	–	–	–	–	–	–	
36	Buccal	2.6	3.5	4.8	3.4	Died		
	Cloacal	–	–	2.3	–			
37	Buccal	3.4	2.4	4.9	3.5	3.6	3.3	Killed
	Cloacal	–	–	–	2.2	–	–	
38	Buccal	**3.5**	4.5	3.7	**4.4**	3.3	Died	
	Cloacal	–	–	–	**3.4**	2.8		
39	Buccal	**3.0**	3.8	4.6	**3.5**	3.2	**2.8**	Killed
	Cloacal	–	–	–	**4.1**	2.6	**2.3**	
40	Buccal	**4.8**	5.1	5.1	**5.1**	4.3	Died	
	Cloacal	–	–	–	**3.3**	2.0		

Viral RNA levels in swabs were detected by real-time RT-PCR as described for Table 6A. RNA titres indicated in bold were selected for deep amplicon sequence analysis. (–) indicates RNA titres $< 10^1$ EID_{50} per ml.

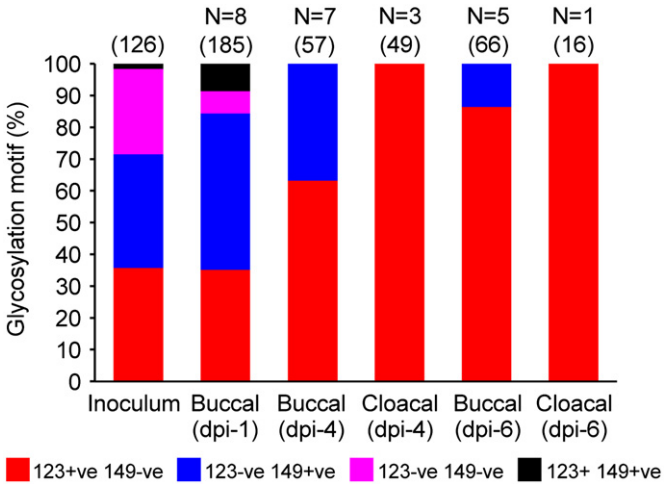

Fig. 3. Changes in the glycosylation motifs in the HA1 polypeptide during infection of turkeys with $10^{3.8}$ and $10^{7.4}$ EID_{50} Italy/1279. The presence (+ve) and absence (-ve) of glycosylation sites at residues 123 and 149 in HA1 are shown as 123+ve and 149-ve (red bars), 123-ve and 149+ve (blue bars), 123-ve and 149-ve (pink bars) and 123+ve and 149+ve (black bars). The bars represent the percentage of sequences encoding a specific pattern of glycosylation at Asn 123 and 149 sites at the indicated time points. N indicates the number of birds sampled at each time point and the number of clones analysed at each time point are indicated in parenthesis.

35), all buccal, encoded no glycosylation at either site. Swabs taken at day 1 pi from this bird showed a mixed population with swabs encoding glycosylation at residue 123 but not 149 (2 clones) and one showing no glycosylation at either site (5 clones). By 6 dpi (82 clones from 5 birds, 5 buccal and 1 cloacal sample) sequences encoding the glycosylation motif at residue 123 of the HA predominated over the cDNAs that encoded a glycosylation motif at residue 149 (Fig. 3). Four of the birds had no sign of heterogeneity with cDNAs prepared from the buccal swab or the cloacal swab encoding a glycosylation sequon at residue 123 but not 149, while only one bird (bird 34) of the five analysed retained a mixed population. At this time point although only three buccal swabs were prepared from bird 35, all now encoded a glycosylation motif at residue 123 but not at residue 149 of HA1 (Fig. S6C).

Transmission of positively selected virus variants to naïve contact animals

The ability of differently glycosylated viruses to be transmitted from infected animals to naïve hosts was also examined. Italy/1279 showed only poor transmission from experimentally infected chickens to naïve contact chickens (data not shown), but transmitted efficiently between turkeys; therefore, the transmission of Italy/1279 between turkeys was studied. A group of 10 turkeys infected with nominal titres of 10^4 EID_{50} were co-housed with 10 naïve turkeys. Virus shedding data revealed that there was a lag time in virus shedding between infected and contact birds (Table 8).

Transmission of virus to all contact naïve birds was observed. RT-PCR positive swab samples from the contact naïve birds were first detected in buccal swabs two to three days after inoculation of the directly challenged birds. cDNAs were prepared from swabs taken from donor animals at day 3 after infection, at which time the majority of contact birds showed signs of infection. cDNAs were prepared from swabs taken from the contact birds were examined at this time and 6 days later (9 days after the infection of the experimentally infected birds) (Table 8). Thus, cDNA clones corresponding to the HA1 coding region were made from swab samples taken at 3 dpi from 8 directly infected birds (921 clones, 8 buccal and 4 cloacal swabs) and 8 contact birds (589 clones, 8 buccal swabs). At day 9 after infection 995 clones were made from 8 buccal and 5 cloacal swabs collected from 8 contact birds. Non-synonymous substitutions compared with the inoculum virus above the level of 1% were observed at nucleotide positions 427, 505, 506 and 800 in samples taken from directly infected

Table 8
Detection of RNA titres in buccal and cloacal swabs collected over 16 days inoculated turkeys with $10^{3.8}$ EID_{50} Italy/1279 virus (white background) and contact naïve turkeys (grey background).

		Days post infection															
Turkey ID	Swab	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
81	Buccal	3.8	2.6	**3.9**	4.2	Died											
	Cloacal	–	–	**3.0**	2.6												
82	Buccal	2.3	4.3	**3.9**	4.8	Died											
	Cloacal	–	–	–	2.1												
83	Buccal	4.1	4.2	**4.1**	Died												
	Cloacal	–	–	**4.3**													
84	Buccal	5.0	3.7	**4.7**	4.2	3.8	2.4	2.3	2.3	1.6	–	–	–	–	–	–	–
	Cloacal	–	–	**2.5**	2.5	2.7	1.6	–	2.0	–	–	–	1.6	–	–	–	–
85	Buccal	–	2.6	2.3	–	2.7	5.0	4.0	2.8	1.6	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	2.1	–	–	2.7	3	–	–	–	–
86	Buccal	4.1	3.1	**3.6**	3.6	4.4	2.2	3.0	2.2	1.6	2.0	2.1	1.8	2.5	2.0	2.3	–
	Cloacal	–	–	–	–	2.4	2.0	2.9	2.0	1.8	1.6	2.6	2.3	2.3	–	–	–
87	Buccal	4.6	3.6	**3.2**	4.7	4.0	2.9	Died									
	Cloacal	–	2.3	**3.2**	3.9	2.5	5.0										
88	Buccal	2.6	3.7	**3.7**	4.2	3.3	3.7	2.3	–	–	–	–	–	–	–	–	–
	Cloacal	–	–	–	–	–	–	2.3	–	1.8	–	–	–	–	–	–	–
89	Buccal	4.0	3.6	4.0	4.8	3.4	2.6	1.6	1.6	–	–	2.1	–	–	–	–	–
	Cloacal	–	–	–	–	1.8	–	1.6	2.6	1.5	–	2.4	–	–	–	–	–
90	Buccal	4.3	3.5	**4.2**	4.2	4.2	2.0	–	2.2	1.5	–	2.1	2.0	–	–	–	–
	Cloacal	–	–	–	–	1.6	–	2.1	1.8	2.3	–	–	–	–	–	–	–
91	Buccal	–	–	**4.2**	5.2	5.0	5.3	3.2	2.1	**2.2**	–	2.3	–	2.3	–	–	–
	Cloacal	–	–	2.0	–	1.6	–	–	–	–	–	–	–	–	–	–	–
92	Buccal	–	3.7	**4.5**	5.5	5.3	4.4	3.6	2.8	1.5	–	–	2.0	–	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	–	––	2.2	–	–	–	–	–
93	Buccal	–	–	**2.1**	2.5	3.2	3.9	3.7	3.6	**4.2**	4.5	3	2.2	2.6	–	–	–
	Cloacal	–	–	–	–	–	–	–	–	2.0	2.0	2.2	–	–	–	–	–
94	Buccal	–	1.8	–	2.9	3.1	5.1	3.3	2.9	**3.5**	4.5	–	Killed				
	Cloacal	–	–	–	–	1.6	–	2.4	3.2	**4.0**	3.3	3.7					
95	Buccal	–	2.5	**4.3**	5.0	4.0	3.1	3.2	2.3	Died							
	Cloacal	–	–	–	–	–	–	–	–								
96	Buccal	–	–	**2.1**	1.8	4.1	5.7	3.2	3.1	**4.1**	4.6	2.5	2.5	–	–	–	–
	Cloacal	–	–	–	1.7	3.1	2.0	–	2.3	**3.1**	3.4	3.2	–	2.2	–	–	–
97	Buccal	–	–	**4.7**	5.7	4.6	5.1	2.7	2.9	**3.2**	3.9	Died					
	Cloacal	–	–	1.6	–	1.7	–	–	2.0	**3.0**	2.0						
98	Buccal	–	–	**2.5**	4.9	3.1	2.1	2.6	3.6	**3.1**	Died						
	Cloacal	–	–	–	–	2.0	–	–	2.0	**2.5**							
99	Buccal	–	–	**2.3**	5.3	2.5	3.2	3.3	2.0	**2.8**	–	–	–	2.1	–	–	-
	Cloacal	–	–	–	–	–	–	–	–	2.0	–	2.5	3.7	–	–	–	–
100	Buccal	–	–	1.6	2.5	3.4	4.4	3.1	3.8	**2.7**	–	2.4	–	2.0	–	–	–
	Cloacal	–	–	1.6	–	–	–	–	–	–	–	–	–	–	–	–	–

Viral RNA levels in swabs were detected by real-time RT-PCR as described for Table 6A. RNA titres indicated in bold were selected for deep amplicon sequence analysis. (–) indicates RNA titres $< 10^1$ EID_{50}/ml.

birds sampled at 3 dpi (Fig. S7B) encoding amino acid substitutions at 125A/T, and 151A/T/D and 249S/N in the HA respectively (Table S2D); and, on the same day, cDNAs prepared from the contact birds showed non-synonymous substitutions at 392, 427 and 800 of the HA gene (Fig. S7C). These nucleotide variants encode the amino acid substitutions at 113M/T, 125A/T and 249S/N respectively (Table S2D). These contact turkeys on 9 dpi showed non-synonymous substitutions at 427, 628, 688, 718 and 800 (Fig. S7D); encoding amino acid substitutions at 125A/T, 192L/M, 212P/I, 222D/N and 249S/N respectively (Table S2D).

All donor animals had virus detectable in buccal swabs at day three, and four also showed cloacal shedding of virus (Table 8). There was a considerable degree of heterogeneity in cDNAs made from swabs taken from the 8 donor animals analysed (Fig. S8). cDNAs from two birds, (84 and 86) exclusively encoded the variant that encoded a glycosylation sequon at 123 but not at 149, with cDNA from bird 84 being prepared from both buccal and cloacal swabs. cDNAs encoding the pattern was seen in the cloacal swab of another bird (bird 81) whilst cDNAs from the buccal swab from this donor bird were mixed, retaining a proportion of about 10% of cDNAs that encoded no glycosylation at either site, and a small proportion that encoded no glycosylation at residue 123 but glycosylation at 149. One bird (bird 83) showed a predominance of cDNAs that encoded no glycosylation at residue 123 but glycosylation at 149 in both buccal and cloacal swabs, mixed with two of the three other possible glycosylation variants as minor species (~10% and less). cDNAs prepared from swab samples taken from bird 87 showed a homogenous population in the cloacal sample which encoded no glycosylation at residue 123 but glycosylation at 149, whilst in the buccal sample the cDNAs encoding the other glycosylation variants were observed with approximately 25% encoding glycosylation at neither residue 123 or 149. cDNAs from one bird (88), in which virus was only detected in buccal swabs, showed a predominance of cDNAs encoding a glycosylation sequon at 123 but not at 149 and cDNAs from another bird (bird 90) the buccal samples produced cDNAs that encoded all four possible variants in proportions not very dissimilar from that of the inoculum. Thus the donor birds had the potential to transmit a variety of differentially glycosylation viruses to the contact birds.

cDNAs prepared form contact turkeys at day 3 or day 9 showed a remarkable lack of heterogeneity in the glycosylation pattern (Fig. S8B, S8C). Of 8 birds analysed at day 3, all from buccal swabs, 6 showed no heterogeneity in glycosylation all encoding a glycosylation sequon at 123 but not at 149, and in the other two birds cDNAs encoding the same pattern of glycosylation predominated with only a small minority (2–5%) showing any alternative pattern, that of encoding no glycosylation at residue 123 but glycosylation at 149. At day 9, none of the cDNAs showed any heterogeneity whether from buccal swabs from all eight birds analysed or from cloacal swabs from 5 birds; all showed a glycosylation sequon at 123 but not at 149.

Discussion

It is recognised that modulation of glycosylation of glycoproteins can have a significant effect on their function. As described above, it is known that the glycosylation status of the HA glycoprotein of influenza viruses can have marked effects on the antigenic properties of the viruses (Munk et al., 1992; Schulze, 1997; Skehel et al., 1984), and on virulence (Matsuoka et al., 2009; Perdue et al., 1995). Importantly, differences in the glycosylation profile have been associated with changes in host range (Banks and Plowright, 2003; Banks et al., 2001; Matrosovich et al., 1999) and notably differences in the glycosylation pattern around the region of the sialic acid receptor binding site of the protein have been observed.

Influenza viruses can vary during the course of infection (Hulse-Post et al., 2005; Palese and Shah, 2007; Taubenberger and Kash, 2010; Webster et al., 1992). To address the possibility of selection of variant viruses that differed in their glycosylation status in the vicinity of the receptor binding site of the HA a series of experimental infections of chickens and turkeys was set up and detailed analysis of HA1 coding regions of the virus HA genes was undertaken following sampling during *in vivo* infection. To examine the proportion of variant viruses in a population rather than the degree of sequence variation in a population, virus genes were studied using a 'deep amplicon sequencing' approach in which numerous cDNA clones of the HA1 coding region of the HA gene were prepared from swab samples taken over the course of infection and each clone was then sequenced. The results presented here indicated that when chickens were infected with viruses with different glycosylation sites around the receptor binding site (N-linked glycosylation at asparagine residues 123 or 149 of the HA, exemplified by Italy/3466 and Italy/4042) no selection of alternative glycosylation sites was observed over the course of infection. However, one of the viruses (Italy/1279), showed heterogeneity within the initial virus population with nucleotide polymorphisms that encoded differing potential

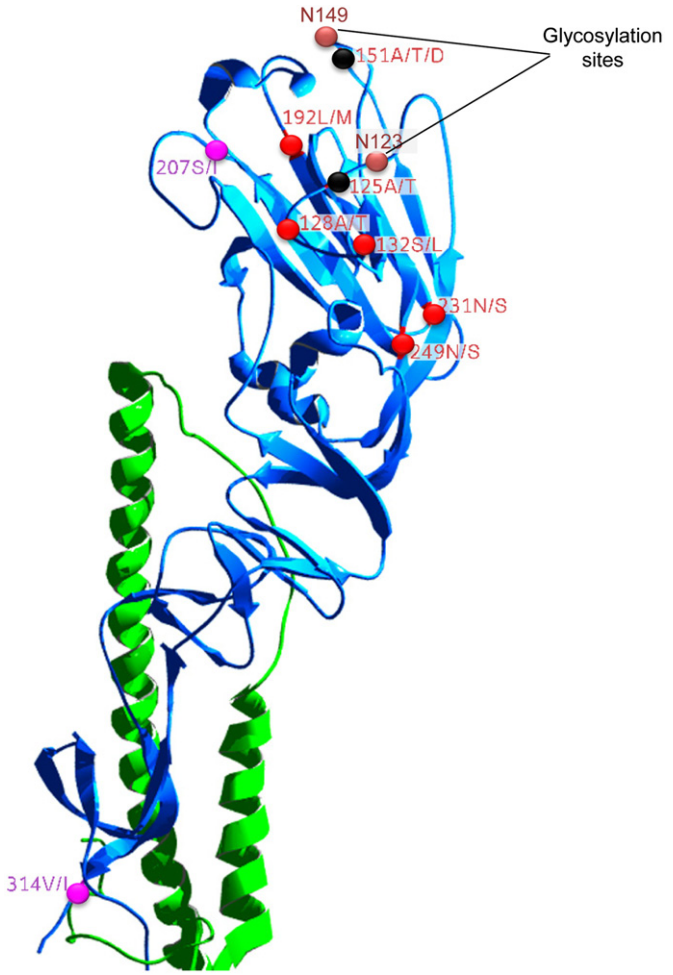

Fig. 4. Schematic representation of monomer of HA glycoprotein (HA1 in blue and HA2 in green) of H7 subtype avian influenza virus PDB number 1T18 (Russell et al., 2004). Amino acids that showed pronounced variation/selection during infection of chickens and turkeys with Italy/1279 and Italy/4042 are indicated in red and purple dots respectively. Substitutions that alter the glycosylation signals are indicated by black dots.

glycosylation patterns in the vicinity of the receptor binding site of the HA. This allowed us to examine possible selection of glycosylation variants within the infected host—in essence this was an *in vivo* virus fitness assay. In both turkeys and chickens, we were able to observe that the mixed population present in the initial inoculum virus resolved within four days after infection. By this time viruses that contained either a glycosylation site at residue 123 or at residue 149 vastly predominated, with glycosylation at 123 prevailing over glycosylation at position 149. It was even more striking that in transmission experiments viruses that encoded a glycosylation motif at 123 represented at least 99% of virus sequences observed. The location of these variable glycosylation sites is shown in Fig. 4.

Viruses with distinct glycosylation patterns might show different tropisms within the infected host with respect to upper gastro/respiratory and lower intestinal locations (and represented by buccal and cloacal swabs). However, in this study we did not observe an exclusive pattern of glycosylation sites in virus RNA recovered from the different swab locations. That said, there was a predominance of viruses with glycosylation in the vicinity of the receptor binding site at residue 123 in cloacal samples taken from either chickens or turkeys and, as the infection progressed over time and in both chickens and turkeys, viruses with the glycosylation motif at resides 123–125 and not at 149–151 predominated (Table S3, S4, Figs. 2 and 3, S4, S6, S8). Whether this observation results directly from differences in tissue tropism is not known. It though is notable that in the samples analysed from turkeys that became infected when in contact with infected birds viruses with the glycosylation motif at resides 123–125 and not at 149–151 massively predominated; in these turkeys only 6 of over 1500 cDNA prepared from buccal and cloacal swabs showed evidence of a potential glycosylation site at 149 and only two clones encoded glycosylation at neither site (Fig. S8). It is striking that this pattern of glycosylation was described in the HPAI viruses evolving from the LPAI virus precursor during the H7N1 epizootic in Italy in 1999–2000 (Banks et al., 2001). Our results reinforce the conclusions drawn by Matrosovich et al. (1999) of the importance of glycosylation in the vicinity of the HA receptor binding site to the adaptation of virus to domestic poultry.

In addition to the selection of viruses *in vivo* that carried substitutions directly affecting the glycosylation signals at Asn 123 and 149 of the HA, a number of minor variants with amino acid substitutions were positively selected during infection of chickens and turkeys (Fig. 4, Table 2). Infection with Italy/1279 resulted in 7 variants at a level of greater than 5% of the total clones analysed, all of which were located in the globular head region of the HA. From Italy/4042 infections, two minor variants emerged; one (residue 207) was a substitution in the globular head region of the HA and the other was located (residue 314) towards the C-terminus of HA1 and outside of the globular head region. No variants were observed at a level of 5% or greater following infection of chickens with Italy/3466. The roles of these changes in the HA are not known.

A similar approach of deep-amplicon sequence analysis has been used to examine the *in vivo* variation of HPAI viruses following experimental infection of ducks, chickens and turkeys. The degree of variation observed *in vivo* showed variants at only low frequency in H7N1 and H5N1 HPAI virus infections of chickens, turkeys and ducks with an overall nucleotide substitution frequency of between 1.69 and 7.04×10^{-4} (Iqbal et al., 2009). In the experiments in the present study a similar overall frequency was observed with the exception of chickens and turkeys infected with Italy/1279, which was approximately 10-fold higher for non-synonymous changes (Table S5). The higher frequency of substitution might have been the result of Italy/1279 having an increased rate of mutation, influenza viruses with an increased rate of mutation having been described previously and have an apparent mutation rate 10-fold higher than the wild type virus (Ding et al., 2010; Suarez-Lopez and Ortin, 1994; Suarez et al., 1992). Recalculation of the nucleotide substitution frequency was carried out for Italy/1279 discounting the three most polymorphic nucleotide positions (427, 505 and 506) that encode the different glycosylation sites of the HA at residues 123 and 149. This reduced the overall HA1 gene mutation frequency, such that Italy/1279 showed a similar nucleotide substitution frequency to those calculated for the other viruses and so it seems unlikely that Italy/1279 carried a polymerase complex with an increased rate of mutation (Table S5).

In this study we have focussed on the variation in glycosylation on the H7 HA as residues 123 and 149 (133 and 158 in H3 numbering). An examination of H7 HA sequences from public databases from the outbreak of avian influenza in the Netherlands in 2003, variation at site 123 was seen in isolates from humans and chickens (Fouchier et al., 2004). Although the number of isolates was low there was no evidence of glycosylation at residue 149. In the introduction we pointed out a similar variation in the glycosylation in the vicinity of the receptor binding site of the H5 HA. In the H5 subtype these residues are 154 and 165 (158 and 169 in H3 numbering). It is noteworthy that the loss of a potential glycosylation site at residue 154 (158 in H3 numbering) is one of a few key adaptations acquired by H5N1 viruses able to transmit between ferrets by the aerosol route (Herfst et al., 2012; Imai et al., 2012).

Conclusion

Our data indicate that when two or more H7N1 viruses that differ in their glycosylation profile around the receptor binding site of the HA are subject to selection *in vivo*, variants can be rapidly selected with a specific pattern of glycosylation within the infected host. In the two species of poultry examined, virus variants having glycosylation at Asn 123 and absent at Asn 149 are predominant in each species. Thus glycosylation near the receptor binding site appears to confer a replicative advantage to these viruses in the infected host as part of a mixed virus population and this virus also effectively transmits to naïve turkeys in contact with experimentally infected turkeys. We submit that our experiments demonstrate that acquisition of glycosylation near the receptor binding site of HA confers a replicative advantage to these viruses in the infected poultry host.

Materials and methods

Virus stocks. LPAI H7N1 virus strains Italy/3466 (accession no. AF364155, sample collected on 1999-9-20), Italy/4042 (accession no. AF364161, collected on 1999-10-23) and Italy/1279 (accession no. AF364142, collected on 1999-4-9) propagated twice only in 9-day-old specific pathogen free (SPF) embryonated fowls' eggs (Charles River, USA) following receipt at the OIE, FAO and EU Reference Laboratory for Avian Influenza and viruses sent to the Reference Laboratory are usually at very low passage, but this passage history is not known for certain prior to receipt. Virus titres were assessed as median egg infectious doses EID_{50}.

Infection of chickens and turkeys. All animal experiments and procedures using live virus were carried out in Defra approved SAPO4/ACDP3 biosafety level 3+ (BSL 3+) facilities at AHVLA-Weybridge. Pre-inoculation blood tests, together with buccal and cloacal swabs were taken to ensure all birds were free of influenza A infection prior to the start of the experiment. Birds, SPF white leghorn chickens (Charles River, USA) or commercial turkeys

from high health status flock (Cobb Breeders, UK), at three weeks of age were inoculated in groups of 10 with nominal doses 10^2 (actual dose $10^{1.8-2.2}$), 10^4 (actual dose $10^{3.8}$) or 10^6 (actual dose $10^{5.4-107.4}$) EID_{50} of virus delivered intranasally in a volume of 0.1 ml. The actual dose of virus used was determined by titration of the dilution of the inoculum for each of the doses used for infection with Italy/3466 and Italy/1279 and for the 10^6 inoculation dose for Italy/4042.

To assess the transmission of the virus from infected turkeys to naïve uninfected contact turkeys, group of 10 three-week-old turkeys were inoculated with 10^2 (actual dose $10^{2.1}$) or 10^4 (actual dose $10^{3.8}$) EID_{50} dose of Italy/1279. Immediately after inoculation, each virus inoculated group was mixed with 10 naïve uninfected turkeys as contacts. All birds were monitored twice daily for clinical signs, and buccal and cloacal swabs were taken daily. Any bird deemed unable to reach food or water, or unduly ill, was killed humanely and recorded as mortality for that day. Each experiment lasted for a maximum of three weeks and any birds surviving at that time were killed humanely.

Viral RNA isolation and determination of viral titres in buccal and cloacal swabs. Buccal and cloacal swabs were taken and the presence of virus RNA detected by an influenza matrix gene one-step quantitative RT-PCR as described (European Union, 2006; Londt et al., 2008). The threshold cycle (*Ct*) value were extrapolated as relative equivalent unit (REU) of viral infectivity titre in $\log_{10} EID_{50}$/ml; based on standard curve constructed from 10-fold serial dilutions of RNA extracted from a known concentration ($\sim 10^6$ EID_{50}) of infective allantoic fluid from the virus inoculum stocks (Italy/3466, Italy 4042 or Italy/1279).

Nucleotide sequencing. An optimised two-step RT-PCR procedure and high-fidelity thermostable polymerase was used to produce cDNA amplicons from aliquots of the virus inocula, and buccal and cloacal swabs. The RT reactions were performed using an influenza virus universal oligonucleotide primer, 5′AGCAAAA GCAGG-3′ with the VersoTM cDNA Kit (Thermo Scientific) according to the supplier's instructions. For PCR amplification each RT reaction (5 μl) was supplemented with 10x buffer, 250 μM each dNTP, 0.2 μM forward and reverse primers in a 50 μl final reaction volume and 1 μl PfuUltra™ II fusion HS DNA polymerase was added (Stratagene, enzyme units were not defined by the manufacturer). The amplification was performed using an initial denaturation step (95 °C for 1 min), followed by 35 cycles of amplification (95 °C for 20 s, 55 °C for 20 s, 72 °C for 45 s) and a final extension (72 °C for 3 min). The PCR products were generally purified by gel electrophoresis and recovered using a gel extraction kit (QIA quick, Qiagen); the gene products were ligated into the pCR-Blunt vector (Invitrogen) and transformed into TOP10 E. coli (Invitrogen). In some cases, if the PCR products were seen as a single band of the gene product, no gel purification step was performed. Positive bacterial colonies containing gene inserts were selected following PCR analysis. Colonies containing plasmids shown to contain inserts were sequenced commercially (GATC Biotech, Constance, Germany).

The viruses used as inocula were subjected to full genome sequence analysis. To generate a consensus a minimum of five independent cDNA clones for each gene segment from each virus were sequenced. The PB2, PB1 and PA genes were amplified in two overlapping amplicons. Primer sequences for PCR amplification, bacterial colony PCR analysis and for deep amplicon sequencing can be made available on request. Sequence data were analysed using the Staden package (pregap4 v1.5 and gap4 v1.0). The full length nucleotide sequences of viruses used as inocula have been submitted to the GenBank database and are available under accession numbers (CY099594 to CY099617). For the HA genes (9–1152 bp) cDNA samples prepared from inoculum viruses, buccal and cloacal swabs, and between 25 and 200 clones per sample were sequenced to detect variant sequences within the amplicon. The selection of swabs for sequence analysis was made on the basis of the titres of viral RNA present in each sample; selected swabs had an estimated RNA level equivalent to $\geq 1 \times 10^2$ EID_{50}/ml of virus.

Analysis of synonymous and non-synonymous substitutions in the HA gene. Sequences were aligned to a reference sequence (the appropriate inoculum consensus sequence, derived from a minimum of 60 cDNAs) using the Gap4 application from the Staden Package (Staden, 1996). The Gap4 report mutation function was then used to generate a report of all base changes. Variation at a frequency of less than 1% was considered as background variation and has not been shown in the graphs.

Analysis of glycosylation status in the HA. Reassortant viruses containing the HA and NA gene of Italy/1279 and internal gene segments of A/PR/8/34 (H1N1) were generated by RG as described previously (Hoffmann et al., 2002). The HA gene harbouring differing glycosylation motifs at position 123 and 149 (described above) were derived from cDNA clones prepared from inoculum stocks of Italy/1279. Differing glycosylation motifs in the HA of the RG viruses were confirmed by sequence analysis as described above. The RG viruses were grown in eggs, purified using continuous sucrose gradient (15–40% *w*/*v*) ultracentrifugation. Deglycosylation of HA1 of purified RG viruses were performed with PNGase F as per the manufacturer's instructions (New England Biolabs). Briefly, approximately 10 μg of each purified virus was mixed with 1x denaturing solution (0.5% SDS and 40 mM DTT) in 10 μl reaction mixture and incubated for 10 min at 100 °C. The reaction volume was then increased to 20 μl by adding 1% NP40, 1x G7 buffer and 2 μl PNGase F and further incubated for 24 h at 37 °C. The viral proteins were separated on SDS-PAGE (any kDa™ mini-protein gel, Bio-Rad) under reducing conditions and stained with Coomassie blue R250.

Acknowledgments

We thank Kolli Reddy and Michelle Hill (IAH) for providing day to day assistance in the laboratory, Linzy Outtrim and Rebecca Gardner for substantial laboratory support (AHVLA), Anna Ramsden for technical preparation (AHVLA) and the AHVLA-ASU team for husbandry and procedural assistance; Andrew Worry (BITS) for providing assistance in bioinformatics analysis; Dr Sue Baigent, Dr Helena Browne, Dr Dave Cavanagh, Dr Steve Wharton and Dr. Dennis Alexander for reviewing the manuscript. This work was supported by funding from the Biotechnology and Biological Sciences Research Council (award nos. BBS/B/00093, BB/E010849/1 and BB/E010806) and the Department for Agriculture and Rural Affairs (projects SE0776 and FT5094).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at 10.1016/j.virol.2012.08.001.

References

Alexander, D.J., 2007. An overview of the epidemiology of avian influenza. Vaccine 25, 5637–5644.

Alexander, D.J., Brown, I.H., 2000. Recent zoonoses caused by influenza A viruses. Revue Scientifique et Technique 19, 197–225.

Baigent, S.J., McCauley, J.W., 2001. Glycosylation of haemagglutinin and stalk-length of neuraminidase combine to regulate the growth of avian influenza viruses in tissue culture. Virus Res. 79, 177–185.

Baigent, S.J., McCauley, J.W., 2003. Influenza type A in humans, mammals and birds: determinants of virus virulence, host-range and interspecies transmission. Bioessays 25, 657–671.
Banks, J., Plowright, L., 2003. Additional glycosylation at the receptor binding site of the hemagglutinin (HA) for H5 and H7 viruses may be an adaptation to poultry hosts, but does it influence pathogenicity? Avian Dis. 47, 942–950.
Banks, J., Speidel, E.S., Moore, E., Plowright, L., Piccirillo, A., Capua, I., Cordioli, P., Fioretti, A., Alexander, D.J., 2001. Changes in the haemagglutinin and the neuraminidase genes prior to the emergence of highly pathogenic H7N1 avian influenza viruses in Italy. Arch. Virol. 146, 963–973.
Brown, I.H., Banks, J., Manvell, R.J., Essen, S.C., Shell, W., Slomka, M., Londt, B., Alexander, D.J., 2006. Recent epidemiology and ecology of influenza A viruses in avian species in Europe and the Middle East. Dev. Biol. (Basel) 124, 45–50.
Butt, K.M., Smith, G.J., Chen, H., Zhang, L.J., Leung, Y.H., Xu, K.M., Lim, W., Webster, R.G., Yuen, K.Y., Peiris, J.S., Guan, Y., 2005. Human infection with an avian H9N2 influenza A virus in Hong Kong in 2003. J. Clin. Microbiol. 43, 5760–5767.
Capua, I., Alexander, D.J., 2009. Avian Influenza and Newcastle Disease: A Field and Laboratory Guide. Springer, Milan; New York.
de Wit, E., Munster, V.J., van Riel, D., Beyer, W.E., Rimmelzwaan, G.F., Kuiken, T., Osterhaus, A.D., Fouchier, R.A., 2010. Molecular determinants of adaptation of highly pathogenic avian influenza H7N7 viruses to efficient replication in the human host. J. Virol. 84, 1597–1606.
Deshpande, K.L., Fried, V.A., Ando, M., Webster, R.G., 1987. Glycosylation affects cleavage of an H5N2 influenza virus hemagglutinin and regulates virulence. Proc. Natl. Acad. Sci. USA 84, 36–40.
Ding, X., Jiang, L., Ke, C., Yang, Z., Lei, C., Cao, K., Xu, J., Xu, L., Yang, X., Zhang, Y., Huang, P., Huang, W., Zhu, X., He, Z., Liu, L., Li, J., Yuan, J., Wu, J., Tang, X., Li, M., 2010. Amino acid sequence analysis and identification of mutations under positive selection in hemagglutinin of 2009 influenza A (H1N1) isolates. Virus Genes 41, 329–340.
European Union, 2006. Diagnostic manual for avian influenza. 〈http://eur-lex.europa.eu/LexUriServ/site/en/oj/2006/l_237/l_23720060831en00010027.pdf〉 (accessed 28 June 2012).
Fouchier, R.A., Schneeberger, P.M., Rozendaal, F.W., Broekman, J.M., Kemink, S.A., Munster, V., Kuiken, T., Rimmelzwaan, G.F., Schutten, M., Van Doornum, G.J., Koch, G., Bosman, A., Koopmans, M., Osterhaus, A.D., 2004. Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. Proc. Natl. Acad. Sci. USA 101, 1356–1361.
Gulati, S., Smith, D.F., Air, G.M., 2009. Deletions of neuraminidase and resistance to oseltamivir may be a consequence of restricted receptor specificity in recent H3N2 influenza viruses. Virol. J. 6, 22.
Guo, Y., Li, J., Cheng, X., 1999. Discovery of men infected by avian influenza A (H9N2) virus. Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi 13, 105–108.
Hatta, M., Gao, P., Halfmann, P., Kawaoka, Y., 2001. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. Science 293, 1840–1842.
Herfst, S., Schrauwen, E.J., Linster, M., Chutinimitkul, S., de Wit, E., Munster, V.J., Sorrell, E.M., Bestebroer, T.M., Burke, D.F., Smith, D.J., Rimmelzwaan, G.F., Osterhaus, A.D., Fouchier, R.A., 2012. Airborne transmission of influenza A/H5N1 virus between ferrets. Science 336, 1534–1541.
Hoffmann, E., Krauss, S., Perez, D., Webby, R., Webster, R.G., 2002. Eight-plasmid system for rapid generation of influenza virus vaccines. Vaccine 20, 3165–3170.
Hulse-Post, D.J., Sturm-Ramirez, K.M., Humberd, J., Seiler, P., Govorkova, E.A., Krauss, S., Scholtissek, C., Puthavathana, P., Buranathai, C., Nguyen, T.D., Long, H.T., Naipospos, T.S., Chen, H., Ellis, T.M., Guan, Y., Peiris, J.S., Webster, R.G., 2005. Role of domestic ducks in the propagation and biological evolution of highly pathogenic H5N1 influenza viruses in Asia. Proc. Natl. Acad. Sci. USA 102, 10682–10687.
Imai, M., Watanabe, T., Hatta, M., Das, S.C., Ozawa, M., Shinya, K., Zhong, G., Hanson, A., Katsura, H., Watanabe, S., Li, C., Kawakami, E., Yamada, S., Kiso, M., Suzuki, Y., Maher, E.A., Neumann, G., Kawaoka, Y., 2012. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. Nature 486, 420–428.
Iqbal, M., Xiao, H., Baillie, G., Warry, A., Essen, S.C., Londt, B., Brookes, S.M., Brown, I.H., McCauley, J.W., 2009. Within-host variation of avian influenza viruses. Philos. Trans. R. Soc. Lon. B Biol. Sci. 364, 2739–2747.
Klenk, H.D., Wagner, R., Heuer, D., Wolff, T., 2002. Importance of hemagglutinin glycosylation for the biological functions of influenza virus. Virus Res. 82, 73–75.
Koopmans, M., Wilbrink, B., Conyn, M., Natrop, G., van der Nat, H., Vennema, H., Meijer, A., van Steenbergen, J., Fouchier, R., Osterhaus, A., Bosman, A., 2004. Transmission of H7N7 avian influenza A virus to human beings during a large outbreak in commercial poultry farms in the Netherlands. Lancet 363, 587–593.
Londt, B.Z., Nunez, A., Banks, J., Nili, H., Johnson, L.K., Alexander, D.J., 2008. Pathogenesis of highly pathogenic avian influenza A/turkey/Turkey/1/2005 H5N1 in Pekin ducks (Anas platyrhynchos) infected experimentally. Avian Pathol. 37, 619–627.
Marinina, V.P., Gambarian, A.S., Bovin, N.V., Tuzikov, A.B., Shilov, A.A., Sinitsyn, B.V., Matrosovich, M.N., 2003. The effect of losing glycosylation sites near the receptor-binding region on the receptor phenotype of the human influenza virus H1N1. Mol. Biol. (Mosk) 37, 550–555.
Matrosovich, M., Zhou, N., Kawaoka, Y., Webster, R., 1999. The surface glycoproteins of H5 influenza viruses isolated from humans, chickens, and wild aquatic birds have distinguishable properties. J. Virol. 73, 1146–1155.
Matsuoka, Y., Swayne, D.E., Thomas, C., Rameix-Welti, M.A., Naffakh, N., Warnes, C., Altholtz, M., Donis, R., Subbarao, K., 2009. Neuraminidase stalk length and additional glycosylation of the hemagglutinin influence the virulence of influenza H5N1 viruses for mice. J Virol. 83, 4704–4708.
Mitnaul, L.J., Matrosovich, M.N., Castrucci, M.R., Tuzikov, A.B., Bovin, N.V., Kobasa, D., Kawaoka, Y., 2000. Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. J. Virol. 74, 6015–6020.
Munier, S., Larcher, T., Cormier-Aline, F., Soubieux, D., Su, B., Guigand, L., Labrosse, B., Cherel, Y., Quere, P., Marc, D., Naffakh, N., 2010. A genetically engineered waterfowl influenza virus with a deletion in the stalk of the neuraminidase has increased virulence for chickens. J. Virol. 84, 940–952.
Munk, K., Pritzer, E., Kretzschmar, E., Gutte, B., Garten, W., Klenk, H.D., 1992. Carbohydrate masking of an antigenic epitope of influenza virus haemagglutinin independent of oligosaccharide size. Glycobiology 2, 233–240.
Olsen, B., Munster, V.J., Wallensten, A., Waldenstrom, J., Osterhaus, A.D., Fouchier, R.A., 2006. Global patterns of influenza a virus in wild birds. Science 312, 384–388.
Palese, P., Shah, M., 2007. Orthomyxoviridae: the viruses and their replication. In: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology. Lippincott, Williams & Wilkins, Philadelphia, pp. 1647–1690.
Perdue, M.L., Latimer, J.W., Crawford, J.M., 1995. A novel carbohydrate addition site on the hemagglutinin protein of a highly pathogenic H7 subtype avian influenza virus. Virology 213, 276–281.
Reading, P.C., Pickett, D.L., Tate, M.D., Whitney, P.G., Job, E.R., Brooks, A.G., 2009. Loss of a single N-linked glycan from the hemagglutinin of influenza virus is associated with resistance to collectins and increased virulence in mice. Respir. Res. 10, 117.
Roberts, P.C., Garten, W., Klenk, H.D., 1993. Role of conserved glycosylation sites in maturation and transport of influenza A virus hemagglutinin. J. Virol. 67, 3048–3060.
Russell, R.J., Gamblin, S.J., Haire, L.F., Stevens, D.J., Xiao, B., Ha, Y., Skehel, J.J., 2004. H1 and H7 influenza haemagglutinin structures extend a structural classification of haemagglutinin subtypes. Virology 325, 287–296.
Schulze, I.T., 1997. Effects of glycosylation on the properties and functions of influenza virus hemagglutinin. J. Infect. Dis. 176 (Suppl 1), S24–28.
Skehel, J.J., Stevens, D.J., Daniels, R.S., Douglas, A.R., Knossow, M., Wilson, I.A., Wiley, D.C., 1984. A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody. Proc. Natl. Acad. Sci. USA 81, 1779–1783.
Sorrell, E.M., Song, H., Pena, L., Perez, D.R., 2010. A 27-amino-acid deletion in the neuraminidase stalk supports replication of an avian H2N2 influenza A virus in the respiratory tract of chickens. J. Virol. 84, 11831–11840.
Staden, R., 1996. The Staden sequence analysis package. Mol. Biotechnol. 5, 233–241.
Suarez-Lopez, P., Ortin, J., 1994. An estimation of the nucleotide substitution rate at defined positions in the influenza virus haemagglutinin gene. J. Gen. Virol. 75, 389–393.
Suarez, P., Valcarcel, J., Ortin, J., 1992. Heterogeneity of the mutation rates of influenza A viruses: isolation of mutator mutants. J. Virol. 66, 2491–2494.
Taubenberger, J.K., Kash, J.C., 2010. Influenza virus evolution, host adaptation, and pandemic formation. Cell Host Microbe 7, 440–451.
Tong, S.X., Li, Y., Rivailler, P., Conrardy, C., Castillo, D.A.A., Chen, L.M., Recuenco, S., Ellison, J.A., Davis, C.T., York, I.A., Turmelle, A.S., Moran, D., Rogers, S., Shi, M., Tao, Y., Weil, M.R., Tang, K., Rowe, L.A., Sammons, S., Xu, X.Y., Frace, M., Lindblade, K.A., Cox, N.J., Anderson, L.J., Rupprecht, C.E., Donis, R.O., 2012. A distinct lineage of influenza A virus from bats. Proc. Natl. Acad. Sci. USA 109, 4269–4274.
Wagner, R., Matrosovich, M., Klenk, H.-D., 2002. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. Rev. Med. Virol. 12, 159–166.
Wagner, R., Wolff, T., Herwig, A., Pleschka, S., Klenk, H.D., 2000. Interdependence of hemagglutinin glycosylation and neuraminidase as regulators of influenza virus growth: a study by reverse genetics. J. Virol. 74, 6316–6323.
Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., Kawaoka, Y., 1992. Evolution and ecology of influenza A viruses. Microbiol. Rev. 56, 152–179.
Webster, R.G., Shortridge, K.F., Kawaoka, Y., 1997. Influenza: interspecies transmission and emergence of new pandemics. FEMS Immunol. Med. Microbiol. 18, 275–279.
Wu, R., Zhang, H., Yang, K., Liang, W., Xiong, Z., Liu, Z., Yang, X., Shao, H., Zheng, X., Chen, M., Xu, D., 2009. Multiple amino acid substitutions are involved in the adaptation of H9N2 avian influenza virus to mice. Vet. Microbiol. 138, 85–91.
Xu, K.M., Li, K.S., Smith, G.J., Li, J.W., Tai, H., Zhang, J.X., Webster, R.G., Peiris, J.S., Chen, H., Guan, Y., 2007. Evolution and molecular epidemiology of H9N2 influenza A viruses from quail in southern China, 2000 to 2005. J. Virol. 81, 2635–2645.
Yen, H.L., Aldridge, J.R., Boon, A.C., Ilyushina, N.A., Salomon, R., Hulse-Post, D.J., Marjuki, H., Franks, J., Boltz, D.A., Bush, D., Lipatov, A.S., Webby, R.J., Rehg, J.E., Webster, R.G., 2009. Changes in H5N1 influenza virus hemagglutinin receptor binding domain affect systemic spread. Proc. Natl. Acad. Sci. USA 106, 286–291.